

Distributed Detection of a Random Process over a Multiple Access Channel under Energy and Bandwidth Constraints

Juan Augusto Maya, *Member, IEEE*, Leonardo Rey Vega, *Member, IEEE*, and Cecilia G. Galarza

Abstract—We analyze a binary hypothesis testing problem built on a wireless sensor network (WSN) for detecting a stationary random process distributed both in space and time with circularly-symmetric complex Gaussian distribution under the Neyman-Pearson framework. Using an analog scheme, the sensors transmit different linear combinations of their measurements through a multiple access channel (MAC) to reach the fusion center (FC), whose task is to decide whether the process is present or not. Considering an energy constraint on each node transmission and a limited amount of channel uses, we compute the miss error exponent of the proposed scheme using Large Deviation Theory (LDT) and show that the proposed strategy is asymptotically optimal (when the number of sensors approaches to infinity) among linear orthogonal schemes. We also show that the proposed scheme obtains significant energy saving in the low signal-to-noise ratio regime, which is the typical scenario of WSNs. Finally, a Monte Carlo simulation of a 2-dimensional process in space validates the analytical results.

Index Terms—Distributed detection; energy and bandwidth constraints; multiple access channel; error exponent; wireless sensor networks;

I. INTRODUCTION

DISTRIBUTED detection based on wireless sensor networks (WSN) is a topic which has attracted great interest in recent years (see [2] and references therein). A typical WSN has a large number of sensor nodes which are generally low-cost battery-powered devices with limited sensing, computing, and communication capabilities. Sensors acquire noisy measurements, perform simple data processing and propagate the information into the WSN to reach a decision about a physical phenomenon happening in the coverage area. The data processing at each node and the propagation of the information within the network are key aspects to be considered to achieve the required performance level.

Network resources, such as energy and bandwidth, are scarce and expensive, but they are key variables when the design is focused on processing latency and detection performance. Clever data processing strategies are required to maximize performance under resources constraints. These constraints could be imposed on a node by node basis, or on the overall network. On any matter, appropriate choices of

the processing strategy could largely impact on the total cost or on the life cycle of the WSN, and as such, habilitate its deployment on remote locations or not.

A. Related work

Distributed detection theory has been much studied in the past. Starting with the seminal work of Tenney and Sandell [3], several results have been derived on how each node compiles the available information and communicates with the *fusion center* (FC) where the final decision on the true state of nature is taken. Under this setup, digital transmission schemes, where appropriate quantization rules has to be designed has been dealt with [4]–[6] (see also references therein). On the other hand, analog communication schemes were also studied in the past (see for example [7]–[9]). For Gaussian networks, with independent and identically distributed (i.i.d.) Gaussian measurements [10], [11], it is known that a simple analog scheme as the scaling and transmission of the noisy measurement, is an optimal joint source-channel scheme in terms of quadratic distortion with power constraint in the sensors. Clearly, this can be viewed as a strong motivation, from a theoretical point of view, for further study of analog schemes for distributed detection problems.

The specific communication strategy from the sensor nodes to the FC has also been extensively studied. The simplest approach is to consider that sensors communicate with the FC using orthogonal parallel channels, like time division multiple access (TDMA) or frequency division multiple access (FDMA) [7], [12]. Clearly, this could not be efficient for large-scale wireless sensor networks where a large bandwidth is required for simultaneous transmissions or a large detection delay is necessary if sensors use the same bandwidth and transmit in different time slots.

A more sophisticated approach is implemented with a multiple access channel (MAC) where sensors transmit simultaneously. In this case, the bandwidth requirement does not necessarily depend on the number of sensors. Transmission over a MAC is appealing because, for certain distributed detection problems, linear mixing of the sensor measurements, naturally performed in a MAC, could be efficiently exploited at the FC. For example, in the case of Gaussian measurements and analog transmissions, this coherent mixing could provide beamforming gains with significant impact on the performance and the energy consumed by the network. The use of a

The three authors are with the University of Buenos Aires (UBA), Buenos Aires, Argentina. L. Rey Vega and C. Galarza are also with the CSC-CONICET, Buenos Aires, Argentina. Emails: {jmaya, lrey, cgalar}@fi.uba.ar

This work is an extension of the results previously presented in [1] and was partially supported by the Peruihl grant (UBA) and the project UBACYT 2002010200250.

MAC channel for the problem of distributed detection has been studied in the past. In [13] linear mixing through a MAC is used for collaboratively computing the network-wide *type* of the measurements (which are assumed to be i.i.d. given the state of nature) taken at each node. As the noise in the channel is asymptotically harmless when the number of sensors grows to infinity, and the type of a set of i.i.d. measurements is a *sufficient statistic*, this strategy has the same optimal performance as the best *centralized* scheme (see also [14], [15]). The specific case in which the possible states of nature are determined by the presence or not of a deterministic signal was considered previously [16]–[18].

However, when the possible states of nature are the presence or absence of a random process, and the nodes use a MAC to communicate with the FC, the problem is more delicate. When the random process is i.i.d. across time and along the sensor nodes, the analog transmission of the log-likelihood ratio (LLR) of the measurement at each node is optimal in the sense that the asymptotic performance converges to that of the centralized detector. This is a consequence of the fact that the network-wide global LLR is simply the sum of the marginal LLRs computed at each node. However, when the random process to be detected is correlated in space and/or time the situation is not so simple. In this case, the measurements taken at each node are not i.i.d. across time and space and the sum of the marginal LLRs (achieved through the coherent combining through the MAC) is clearly suboptimal, even when there is no fading [19]. As the network-wide LLR is in general a non-linear function of the marginal LLRs, linear combining of the MAC cannot provide the optimal statistic to the FC, as in the case of i.i.d. data. This is also evident given that when each node sends only its marginal LLR, it is neglecting the correlation with measurements of other nodes.

Problems with correlated observations could be considerably challenging [20]. Design of distributed processing strategies to benefit from the correlation among data is, in general, an open problem. It is well known that signal correlation can help to improve the detection performance, specially when low quality sensor measurements are available [9], [21]. Clearly, a clever use of the correlation in space and/or time requires of cooperation among nodes. This problem has also been studied in previous works [22], [23]. However, the specific case of the distributed detection of a random process where the nodes communicate with the FC through a MAC still deserves some more study.

B. Contributions

We will consider a distributed detection scenario where all the sensors communicate with the FC through a MAC. Each sensor obtains a local measurement that is a realization of a Gaussian stochastic process arbitrarily correlated in space and possibly in time. Our goal will be to analyze the asymptotic performance of the detection scheme, when the number of sensors approaches to infinity. In particular, we look at the error exponents under the Neyman-Pearson scenario. We will consider several processing strategies that

exploit correlation in space and time. These strategies can be briefly described as follows. On a first step, each sensor takes a single measurement or a set of measurements. Then, in a synchronous manner, all the sensors transmit different analog linear combinations of their measurements using several MAC uses. Finally, the FC gathers all the data and constructs an appropriate statistic to make the decision. The cooperation among nodes is achieved through multiple channel uses. As these multiple uses are clearly pricey, we also impose that the sensors have a limited energy budget to be spent on these transmissions, and we carefully select the energy used in each channel use in order to comply with this budget. We obtain then the optimal energy allocation policy which depends on the statistical properties of the process to be detected. It is shown however that for a large number of nodes, knowledge of these statistics is needed only at the FC. As a side result, the optimal number of MAC uses is also obtained. It is shown, that in general the required number of channel uses is not very large for general correlated processes. This implies that this form of cooperation does not impose severe penalties in bandwidth or delay, allowing for a close to optimal performance in terms of error exponents. The performance gains obtained with the proposed schemes are more important when the channel signal to noise ratio (SNR_C) is small, which is the usual scenario for WSNs designed for long life cycle.

C. Notation and Organization

Vectors are written in boldface and matrices in capital letters. A_n and B_{nm} are square and rectangular matrices of sizes $n \times n$ and $n \times m$, respectively. $\det A_n$ and $\text{tr} A_n$ are the determinant and the trace of A_n , respectively. $(\cdot)^T$ and $(\cdot)^H$ denote transpose and transpose conjugate. We do not distinguish between random variables and their realization values. $p(x|\mathcal{H})$ is the probability density function (p.d.f.) of x conditioned to \mathcal{H} . $\mathbb{P}_i(\cdot)$ and $\mathbb{E}_i(\cdot)$ are the probability and the expectation, respectively, computed under hypothesis \mathcal{H}_i . The pre-image of the set A through the function $\phi(\cdot)$ is $\phi^{-1}(A) = \{\nu \in \text{dom}(\phi) : \phi(\nu) \in A\}$. $\mathbb{1}(\nu \in A)$ is the indicator function, i.e., it is 1 if $\nu \in A$ and 0 otherwise. The set of absolutely integrable and essentially bounded functions of support $[0, 1]$ are, respectively, $L^1([0, 1]) = \{f : \int_0^1 |f(\nu)| d\nu < \infty\}$ and $L^\infty([0, 1]) = \{f : \text{ess sup } |f(\nu)| < \infty\}$.

The paper is organized as follows. In Section II, we formulate the binary hypothesis testing problem and describe the MAC between the sensors and the FC. In Section III, we present the centralized detector and then we develop two strategies for distributed detection. In Section IV, we enunciate auxiliary tools used to compute the error exponents in Section V. In Section VI, we show numerical results for one-dimensional networks and in Section VII we extend the results for multi-dimensional networks. In Section VIII we apply the results for detecting a 2D spatial random process and in Section IX we elaborate on the main conclusions. Technical proofs are provided in the appendices.

II. DETECTION PROBLEM

We first consider a network where the nodes are distributed along a line and each sensor takes a single measurement. Networks with more dimensions (in space and/or time) will be considered in section VII. The measurement at the k -th sensor under each hypothesis is:

$$\begin{cases} \mathcal{H}_1: & x_k = s_k + v_k, \\ \mathcal{H}_0: & x_k = v_k, \end{cases} \quad k = 1, \dots, n. \quad (1)$$

We assume that s_k is a zero-mean circularly-symmetric complex Gaussian stationary process with variance σ_s^2 and power spectral density (PSD) $\phi(\nu)$, and v_k is a zero-mean circularly-symmetric complex white Gaussian noise independent of s_k with variance σ_v^2 . Thus, x_k is Gaussian distributed either under \mathcal{H}_0 or \mathcal{H}_1 . We define the following column vectors: $\mathbf{s} = [s_1, \dots, s_n]^T$, $\mathbf{v} = [v_1, \dots, v_n]^T$ and $\mathbf{x} = [x_1, \dots, x_n]^T$. The covariance matrix of \mathbf{v} is $\sigma_v^2 I_n$ where I_n is the identity matrix of dimension n . The signal vector \mathbf{s} has a Toeplitz covariance matrix $\Sigma_n \triangleq \Sigma_n(\phi)$ whose (i, j) -th element is completely characterized by $\phi(\nu)$ as,

$$(\Sigma_n)_{i,j} = \int_0^1 \phi(\nu) e^{-j2\pi\nu(i-j)} d\nu \quad 1 \leq i, j \leq n.$$

where ν is the normalized frequency. The covariance matrices of \mathbf{x} are $\Sigma_{0,n} = \sigma_v^2 I_n$ under \mathcal{H}_0 , and $\Sigma_{1,n} = \Sigma_n + \sigma_v^2 I_n$ under \mathcal{H}_1 .

Remark 1. The Toeplitz assumption allows us to manage the correlation between the measurements at the nodes in a simple manner when the number of them grows unbounded. It has also been considered in [9], [17]. Physically, it can be linked to a situation in which the sensor nodes are located on a regular grid and the continuous random process in space is stationary. More general correlation models include Gaussian random fields [22], [23]. However, under the setup considered in this paper, interesting conclusions and closed form results would more involved. For that reason, we will work with processes described by Toeplitz covariance matrices, which allows us to consider sufficiently general correlation models.

We analyze a sensor network where the nodes communicate with the FC through a MAC with equal gain on each link node-FC. This is a simplification for the general setup, however it is a reasonable approximation for networks deployed in rural and remote areas. In this case, each node has a line of sight with the FC, the nodes are steady, and the surroundings do not vary much. Under this scenario, node synchronization and channel inversion at the nodes are feasible. This assumption has been extensively considered in the past [13], [24]–[28].

Consider now n' channel uses¹, with $n' \leq n$. Each channel use is associated to either an orthogonal time slot or a frequency band. Then, the processing strategy may use n' time slots, n' frequency bands or a combination of them such that the product of time slots and frequency bands is n' . During a channel use, sensors communicate with the FC through a noisy MAC without any other interference. The signal collected at

the FC is a noisy version of the coherent superposition of the symbols transmitted by the n sensors through the MAC,

$$z_{k'} = \sum_{k=1}^n g_{kk'}(x_k) + w_{k'}, \quad k' = 1, \dots, n', \quad (2)$$

where $g_{kk'}(\cdot)$ is the encoding function used by the sensor k at channel use k' , and $w_{k'}$ is the zero-mean communication noise with variance σ_w^2 and circularly-symmetric complex Gaussian distribution independent of everything else. An illustration of the distributed scheme is shown in Fig. 1.

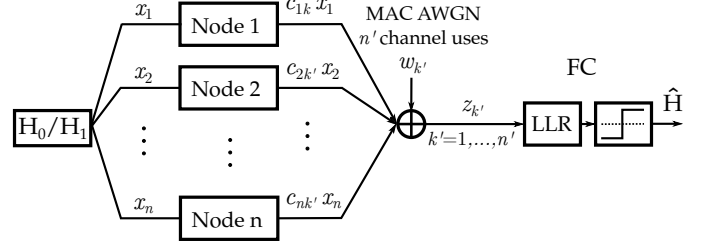


Fig. 1. Decentralized detection scheme for a WSN with linear encoding functions: $g_{kk'}(x_k) = c_{kk'} x_k$.

III. DETECTORS

A. Centralized detector

Suppose that the FC has direct access to the complete measured vector \mathbf{x} through n orthogonal noiseless communication channels. The appropriate use of these measurements allows the FC to construct the optimal centralized detector (CD) [29]. In a distributed setting, with noisy links from the sensors to the FC, this detector provides an upper bound (not necessarily tight) on the performance of any distributed scheme. Consider the Neyman-Pearson problem for a fixed false alarm probability level α , where a false alarm event occurs when \mathcal{H}_1 is declared but \mathcal{H}_0 is true. The associated normalized logarithmic likelihood ratio (LLR) is [29]

$$\begin{aligned} T_c^n(\mathbf{x}) &= \frac{1}{n} \log \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} \\ &= \frac{1}{n} \left(\mathbf{x}^H (\Sigma_{0,n}^{-1} - \Sigma_{1,n}^{-1}) \mathbf{x} - \log \frac{\det \Sigma_{1,n}}{\det \Sigma_{0,n}} \right). \end{aligned} \quad (3)$$

Now, according to the Neyman-Pearson theorem, the optimum centralized test chooses \mathcal{H}_1 if $T_c^n(\mathbf{x}_n) > \tau_n$, and \mathcal{H}_0 otherwise, where the threshold of the test τ_n depends on α .

When the process $\{s_k\}$ is spatially correlated, the covariance matrix $\Sigma_{1,n}$ is not diagonal. Hence, (3) cannot be expressed as the sum of the marginal LLRs from each node, and the mixing property of the MAC does not lead to the global LLR as in [19]. However, it is possible to implement simple distributed detection schemes to reconstruct, in some way, the centralized statistic at the FC through multiple channel uses.

B. Distributed detector

We consider distributed detection (DD) schemes where the nodes make n' channel uses through a MAC. In this work, we

¹We will refer to channel uses or degrees of freedom (DoF) interchangeably.

restrict ourselves to the case where $g_{kk'}(\cdot)$ are linear encoding functions. Thus, in the case of one-dimensional networks, each sensor transmits scaled versions of its local measurement through the MAC and (2) results in:

$$z_{k'} = \mathbf{c}_{k'}^H \mathbf{x} + w_{k'}, \quad k' = 1, \dots, n',$$

where $\mathbf{c}_{k'} = [c_{1k'}, c_{2k'}, \dots, c_{nk'}]^T$. Define the *precoding matrix* as $C_{nn'} = [\mathbf{c}_1, \dots, \mathbf{c}_{n'}]$ and $\mathbf{z} = [z_1, \dots, z_{n'}]^T$. Then

$$\mathbf{z} = C_{nn'}^H \mathbf{x} + \mathbf{w}, \quad (4)$$

where $\mathbf{w} = [w_1, \dots, w_{n'}]^T$ has covariance matrix $\sigma_w^2 I_{n'}$. Considering the Neyman-Pearson problem, the normalized LLR at the FC is:

$$\begin{aligned} T_d^n(\mathbf{z}) &= \frac{1}{n} \log \frac{p(\mathbf{z}|\mathcal{H}_1)}{p(\mathbf{z}|\mathcal{H}_0)} \\ &= \frac{1}{n} \left(\mathbf{z}^H \left(\Xi_{0,n'}^{-1} - \Xi_{1,n'}^{-1} \right) \mathbf{z} - \log \frac{\det \Xi_{1,n'}}{\det \Xi_{0,n'}} \right), \end{aligned} \quad (5)$$

where $\Xi_{1,n'} = C_{nn'}^H \Sigma_n C_{nn'} + \sigma_v^2 C_{nn'}^H C_{nn'} + \sigma_w^2 I_{n'}$, and $\Xi_{0,n'} = \sigma_v^2 C_{nn'}^H C_{nn'} + \sigma_w^2 I_{n'}$ are the covariance matrices of \mathbf{z} under \mathcal{H}_1 and \mathcal{H}_0 , respectively. The false alarm and the miss error probability are defined as $P_{fa}^n = \mathbb{P}_0(T_d^n(\mathbf{z}) > \tau_n)$ and $P_m^n = \mathbb{P}_1(T_d^n(\mathbf{z}) < \tau_n)$, respectively. The average energy consumed by the k -th node during n' transmissions is:

$$E_k = \mathbb{E} \left[\sum_{k'=1}^{n'} |c_{kk'} x_k|^2 \right] = \mathbb{E}[|x_k|^2] \sum_{k'=1}^{n'} |c_{kk'}|^2, \quad k = 1, \dots, n,$$

where $c_{kk'}$ is the (k, k') -th element of $C_{nn'}$, $\mathbb{E}[|x_k|^2] = \sigma_v^2 + p_1 \sigma_s^2$, and p_1 is the *a priori* probability of the state of nature \mathcal{H}_1 . When p_1 is unknown, a natural upper-bound for E_k is obtained by taking $p_1 = 1$.

The problem is then to obtain appropriate precoding matrices $C_{nn'}$ such that the constraints on energy and channel uses are satisfied in the Neyman-Pearson setting:

Problem 1 (Best linear precoding strategy). Consider that n sensors take measurements according to (1) and the communication model is (4). When the per-sensor average energy constraint is E_t , and the probability of false alarm is less than α , the best linear precoding strategy is obtained by solving

$$\inf_{C_{nn'} \in \mathcal{A}_{nn'}} P_m^n,$$

with $\mathcal{A}_{nn'} = \{C_{nn'} \in \mathbb{C}^{n \times n'} : P_{fa}^n \leq \alpha, \frac{1}{n} \text{tr}(C_{nn'} C_{nn'}^H) \leq \frac{E_t}{\sigma_s^2 + \sigma_v^2}\}$.

This problem is not convex in $C_{nn'}$, and closed-form solutions are not readily available. Therefore, we restrict the precoding matrices $C_{nn'}$ to have orthogonal columns and consider WSNs with large number of nodes where it is relevant to compute the *error exponent* of P_m^n when $P_{fa}^n \leq \alpha$. Then, the problem that we attack is as follows:

Problem 2 (Best asymptotic orthogonal precoding strategy). Consider that n sensors take measurements according to (1) and they communicate with the FC through a MAC as in (4). Consider also that $\frac{n'}{n} \rightarrow \beta$ when $n \rightarrow \infty$ for a given

asymptotic fraction of DoF $\beta \in (0, 1]$. When the per-sensor average energy constraint is E_t , and the level of false alarm is limited to α , the best asymptotic orthogonal precoding strategy is obtained by solving

$$\sup_{\{C_{nn'}\}_{n=1}^{\infty} \in \{\mathcal{A}_{nn'}^{\text{orth}}\}_{n=1}^{\infty}} - \lim_{n \rightarrow \infty} \frac{1}{n} \log P_m^n$$

with $\mathcal{A}_{nn'}^{\text{orth}} = \{C_{nn'} \in \mathbb{C}^{n \times n'} : C_{nn'} = V_{nn'} \Delta_{n'}, V_{nn'}^H V_{nn'} = I_{n'}, \Delta_{n'} = \text{diag}(\gamma_1^n, \dots, \gamma_{n'}^n), P_{fa}^n \leq \alpha, \frac{1}{n} \sum_{k'=1}^{n'} (\gamma_{k'}^n)^2 \leq \frac{E_t}{\sigma_s^2 + \sigma_v^2}\}$, where the coefficients $\{\gamma_{k'}^n\}$ control the energy transmitted by the sensors on each channel use.

To tackle this problem, we consider two different parameterizations for the sequence of precoding matrices $\{C_{nn'}\}$. Before elaborating on them, we make the following definition:

Definition 1 (Transmitted Modes Set). Let L be the uniform probability measure defined on the interval $[0, 1]$, and denote by $L[\phi^{-1}(A)]$ the measure of the set $\phi^{-1}(A)$. Define the complementary cumulative distribution function $\Omega(t)$ as the measure of the set of frequencies ν such that $\phi(\nu)$ is at least t , i.e., $\Omega(t) = L[\phi^{-1}((t, +\infty))]$. For $\beta \in (0, 1]$, define $\Theta_\beta := \{\nu \in [0, 1] : \Omega(\phi(\nu)) \leq \beta\}$ as the transmitted modes set of Lebesgue measure β whose elements are the frequencies ν that take on the largest values of $\phi(\nu)$.

Now, to solve Problem 2, we consider two strategies:

Definition 2 (Principal Component Strategy in a MAC channel, PCS-MAC). Let $n \in \mathbb{N}$ and $n' \triangleq n'(n) \leq n$. Assume that $\{\mathbf{u}_k\}_{k=1}^{n'}$ is a basis of eigenvectors for Σ_n and $\{\lambda_k^n\}_{k=1}^{n'}$ the corresponding eigenvalues, where $\lambda_1^n \geq \lambda_2^n \geq \dots \geq \lambda_{n'}^n$. We choose the precoding matrix for the PCS-MAC strategy as

$$C_{nn'} = U_{nn'} \Delta_{n'}, \quad (6)$$

where $U_{nn'} = [\mathbf{u}_1, \dots, \mathbf{u}_{n'}]$, $\Delta_{n'} = \text{diag}(\gamma_1^n, \dots, \gamma_{n'}^n)$ and $n' = \max\{k \in [1, n] : \Omega(\lambda_k^n) \leq \beta\}$.

To implement this strategy, the rows of $C_{nn'}$ are required to be known at the corresponding nodes. For that, either each node should perform a local eigenvalue-eigenvector decomposition, or the FC should communicate the corresponding row of $C_{nn'}$ to each sensor through a feedback channel. The first option entails a more expensive node deployment, and the second option involves the implementation of a high throughput multi-terminal communication link whose complexity would grow linearly as the number of nodes increases. To reduce complexity, we replace the basis of eigenvectors in PCS-MAC by the Fourier basis, and we propose the next strategy:

Definition 3 (Principal Frequencies Strategy in a MAC channel, PFS-MAC). Consider the power spectral density (PSD) $\phi(\nu)$ and let $n \in \mathbb{N}$ and $n' \triangleq n'(n) \leq n$. For each n , let $(j_1, j_2, \dots, j_{n'})$ be a permutation of $\{1, 2, \dots, n\}$ such that $\phi(\frac{j_1-1}{n}) \geq \phi(\frac{j_2-1}{n}) \geq \dots \geq \phi(\frac{j_{n'}-1}{n})$. We choose the precoding matrix PFS-MAC strategy as

$$C_{nn'} = F_{nn'} \Delta_{n'} \quad (7)$$

where $n' = \max \{k \in [1, n] : \Omega(\phi((j_k^n - 1)/n)) \leq \beta\}$, $F_{nn'} = [f_{j_1}, \dots, f_{j_{n'}}]$ is a sub-matrix of the DFT matrix of order n , i.e., $f_{k'} = [f_{1k'}, f_{2k'}, \dots, f_{nk'}]^T$ with $f_{kk'} = \frac{1}{\sqrt{n}} \exp(j2\pi(k-1)(k'-1)/n)$ for $k = 1, \dots, n$, $k' = j_1, \dots, j_{n'}$ and $\Delta_{n'} = \text{diag}(\gamma_{j_1}^n, \dots, \gamma_{j_{n'}}^n)$.

Under this strategy, at each channel use, the FC receives a noisy version of a given frequency bin of the DFT of the measurement vector. For that, the nodes are required to know the index number of the DFT bin to be transmitted only. This information is common to all nodes and it may be broadcasted by the FC on a low rate feedback channel.

Let us interpret the above definitions. For the proposed decentralized strategies, we have imposed that the number of channel uses be limited to $n' \leq n$. Using a MAC in each channel use, we can limit the number of channel uses without discarding any measurement of the sensors and distributing the transmitted energy more efficiently. We will see in Corollary 2 that PCS-MAC and PFS-MAC results in the asymptotic transmission of the components of the spectrum $\phi(\nu)$ in the set Θ_β . Intuitively, we see that this is the best we can do, as the large components of $\phi(\nu)$ are the *more informative* about the state of nature \mathcal{H}_1 , when the fraction of channel uses approaches to β .

Remark 2. For PCS-MAC, $c_{kk'} = u_{kk'} \gamma_{k'}$ where $u_{kk'}$ is the (k, k') -th element of $U_{nn'}$, and the average energy consumed in each node is $E_k = (p_1 \sigma_s^2 + \sigma_v^2) \sum_{k'=1}^{n'} |u_{kk'}|^2 (\gamma_{k'}^n)^2$. This depends on each sensor k . If we additionally consider the average over all the sensors, we obtain the energy constraint considered in Problem 2:

$$\frac{1}{n} \sum_{k'=1}^{n'} (\gamma_{k'}^n)^2 \leq \frac{E_t}{\sigma_s^2 + \sigma_v^2}. \quad (8)$$

On the other hand, in the case of PFS-MAC, $c_{kk'} = \frac{1}{\sqrt{n}} \gamma_{k'}^n e^{j2\pi(k-1)(k'-1)/n}$, and the average energy consumed in each node is $E_k = (p_1 \sigma_s^2 + \sigma_v^2) \frac{1}{n} \sum_{k'=1}^{n'} (\gamma_{k'}^n)^2$. This is independent of the sensor k and therefore (8) is actually an energy constraint on *each* single sensor.

It is convenient to express $\gamma_{k'}^n$ as the sampled version of a function $\xi : [0, 1] \rightarrow \mathbb{R}_+$,

$$\gamma_{k'}^n = \sqrt{\xi\left(\frac{k'-1}{n}\right)}, \quad k' = 1, \dots, n', \quad (9)$$

where $\xi(\cdot)$ is a Riemann integrable function defined as the asymptotic energy profile. The average energy constraint when n goes to infinity results in

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n'} (\gamma_k^n)^2 = \int_{\Theta_\beta} \xi(\nu) d\nu \leq \frac{E_t}{\sigma_s^2 + \sigma_v^2}. \quad (10)$$

Remark 3. A similar general setup was previously considered in [12]. However, the authors imposed certain restrictions to the overall model that were less in our setup. For instance, the communication channels between the nodes and the FC were orthogonal, instead of a MAC as it is analyzed here.

It was also assumed that each sensor measured the same realization of a process under \mathcal{H}_1 disturbed by different noise realizations. That would have been the case of a random process maximally correlated in space. Our analysis allows for general correlation functions both in space and/or time. In addition and more importantly, by introducing the coefficients $\{\gamma_{k'}^n\}$, we optimize the energy profile for a given energy budget and we find the optimum number of MAC uses to achieve the best error exponent among the orthogonal strategies.

IV. PRELIMINARY TOOLS

Definition 4 (Weak and Strong Norms). Let A be a Hermitian $n \times n$ matrix with eigenvalues $\{\lambda_k\}_{k=1}^n$, its weak (normalized Frobenius) and strong (spectral) norms are, respectively,

$$|A| = \left(\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}, \quad \|A\| = \max_{1 \leq k \leq n} |\lambda_k|.$$

Definition 5 (Wiener Class Functions). A function $\phi(\nu)$ defined on the normalized frequency interval $[0, 1]$ is said to be in the Wiener class if it has a Fourier series with absolutely summable Fourier coefficients a_k , i.e., $\sum_{k=-\infty}^{\infty} |a_k| < \infty$ and

$$\phi(\nu) = \sum_{k=-\infty}^{\infty} a_k e^{j2\pi\nu k}, \quad a_k = \int_0^1 \phi(\nu) e^{-j2\pi\nu k} d\nu. \quad (11)$$

Definition 6. The circulant matrix $B_n \triangleq B_n(\phi)$ generated by the samples of the function $\phi(\nu)$ with $\nu = \frac{i-1}{n}$, $i = 1, \dots, n$, is completely specified by its first row $\mathbf{b}_n = [b_1^n, \dots, b_n^n]$,

$$B_n(\phi) = \begin{bmatrix} b_1^n & b_2^n & \dots & b_n^n \\ b_n^n & b_1^n & \ddots & b_{n-1}^n \\ & \ddots & \ddots & \\ b_2^n & b_3^n & \dots & b_1^n \end{bmatrix} \quad (12)$$

where $b_k^n = \sum_{i=1}^n \phi(\frac{i-1}{n}) e^{j\frac{2\pi}{n}(i-1)(k-1)}$, $k = 1, \dots, n$. The eigenvalues of this matrix are given by $\{\phi(\frac{i-1}{n})\}_{i=1}^n$.

Definition 7 (Asymptotically Equivalent Matrices). Two sequences of $n \times n$ matrices $\{A_n\}$ and $\{B_n\}$ are said to be asymptotically equivalent, $A_n \sim B_n$, if

1) A_n and B_n are uniformly bounded in strong norm, i.e.:

$$\|A_n\|, \|B_n\| \leq M < \infty, \quad n = 1, 2, \dots \quad (13)$$

2) $A_n - B_n$ converges to zero in weak norm as $n \rightarrow \infty$, i.e.:

$$\lim_{n \rightarrow \infty} |A_n - B_n| = 0. \quad (14)$$

Lemma 1 (Asymptotically Equivalence of Toeplitz and Circulant Matrices). Let $A_n(\phi) = [a_{i-j}]_{1 \leq i, j \leq n}$ be a Toeplitz matrix with the function $\phi(\nu)$ in the Wiener class related to a_k as in (11) and let $B_n(\phi)$ be a circulant matrix as in (12). Then, $A_n(\phi)$ and $B_n(\phi)$ are asymptotically equivalent.

Proof: See [30, p. 53]. ■

Theorem 1 (Asymptotic Toeplitz Distribution Theorem). Assume that $\phi(\nu)$ is a Wiener class function and Σ_n the

Hermitian Toeplitz matrix related to ϕ . Let $\{\lambda_k^n\}_{k=1}^n$ be the eigenvalues of Σ_n , and $\delta_2 \geq \lambda_1^n \geq \dots \geq \lambda_n^n \geq \delta_1$. If $F(\cdot)$ is a continuous function defined on the interval $[\delta_1, \delta_2]$, such that $\int_{\nu: \phi(\nu)=\delta} F(\phi(\nu)) d\nu = 0$ for any $\delta \in [\delta_1, \delta_2] \cap \Delta$ where $\Delta \subseteq [0, \infty)$, then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n F(\lambda_k^n) \mathbb{1}(\lambda_k^n \in \Delta) = \int_{\phi^{-1}(\Delta)} F(\phi(\nu)) d\nu.$$

Proof: The proof follows from [30, Corollary 4.1]. ■

Definition 8 (Large Deviation Principle, LDP). If G° and \bar{G} are the interior and closure of a set $G \subset \mathbb{R}$, respectively, we say that the sequence of random variables $\{Y_n\}_{n=1}^\infty$ satisfies the LDP with rate function $\Lambda^*(x)$ if, for any $G \subset \mathbb{R}$ we have

$$\begin{aligned} -\inf_{x \in G^\circ} \Lambda^*(x) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Y_n \in G) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(Y_n \in G) \leq -\inf_{x \in \bar{G}} \Lambda^*(x). \end{aligned} \quad (15)$$

The set G is said to satisfy the Λ -continuous property if $\inf_{x \in G^\circ} \Lambda^*(x) = \inf_{x \in \bar{G}} \Lambda^*(x)$. Thus, the lower and upper bounds in (15) coincide, and the exponent ϵ_G is defined as:

$$\epsilon_G = \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(Y_n \in G) = \inf_{x \in G^\circ} \Lambda^*(x) = \inf_{x \in \bar{G}} \Lambda^*(x). \quad (16)$$

The Gärtner-Ellis theorem [31] is a handy result for the computation of a good rate function for a general sequence of random variables. For this result to be true, certain conditions on the asymptotic logarithmic moment-generating function (LMGF), defined as the limit $\Lambda(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda^n(nt)$, with $\Lambda^n(t) = \log \mathbb{E}[e^{tY_n}]$ need to be satisfied. A particular critical condition is the steepness of $\Lambda(t)$ at the boundary of its domain. When Y_n is a sequence of Gaussian quadratic forms constructed from a stationary Gaussian random process (which is the case for $T_d^n(z)$ and $T_c^n(x)$) this steepness condition is a delicate issue [32]. The asymptotic *bad behavior* of some eigenvalues of the corresponding Toeplitz matrices of the stationary process could have a critical role in the behavior of $\Lambda(t)$. For that reason, and for the particular case of the hypothesis testing problem based on likelihood ratio test, we need the following result:

Theorem 2 (Modified Gärtner-Ellis theorem for Gaussian LLR). Let $\{T^n\}_{n=1}^\infty$ be a sequence of LLRs of complex circularly symmetric Gaussian random variables drawn from a stationary Gaussian random process with spectral density $h_i(\nu)$ under \mathcal{H}_i , $i = 0, 1$, and define its LMGF as

$$\Lambda_i(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \Lambda_i^n(nt), \quad \Lambda_i^n(t) = \log \mathbb{E}_i[e^{tT^n}]. \quad (17)$$

Consider the following assumptions:

- A₁) $h_i(\nu)$ is in the Szegö class, i.e., $\log h_i(\nu) \in L^1([0, 1])$, $i = 0, 1$.
- A₂) The ratio of spectral densities is essentially bounded, i.e., $\frac{h_i(\nu)}{h_{\setminus i}(\nu)} \in L^\infty([0, 1])$, where $\setminus i = 1$ if $i = 0$ and $\setminus i = 0$ if $i = 1$.

Then, under \mathcal{H}_i , the sequence $\{T^n\}$ satisfies the LDP in Def. 8 with good rate function $\Lambda_i^*(x)$ computed through the Fenchel-Legendre transform of $\Lambda_i(t)$, $i = 0, 1$:

$$\Lambda_i^*(x) = \sup_{t \in \mathbb{R}} \{xt - \Lambda_i(t)\}, \quad (18)$$

$$\text{with } \Lambda_i(t) = -\int_0^1 \left\{ \log \left(1 + t \frac{h_0(\nu) - h_1(\nu)}{h_{\setminus i}(\nu)} \right) + t \log \frac{h_1(\nu)}{h_0(\nu)} \right\} d\nu.$$

Proof: See [32, Proposition 7]. ■

V. ERROR EXPONENTS

We are now ready to compute the error exponents for both the centralized detector and the decentralized detector considering both strategies, PCS-MAC and PFS-MAC. Before proceeding, we formulate some handy definitions. Consider a spectral density $\Gamma \triangleq \Gamma(\nu)$, a set of frequencies Θ , and a threshold τ . Then, we define the following functionals:

$$m_0(\Gamma) = \int_\Theta \left\{ \frac{\Gamma(\nu)}{1 + \Gamma(\nu)} - \log(1 + \Gamma(\nu)) \right\} d\nu, \quad (19)$$

$$m_1(\Gamma) = \int_\Theta \{ \Gamma(\nu) - \log(1 + \Gamma(\nu)) \} d\nu. \quad (20)$$

$$\begin{aligned} \kappa_{fa}(\Gamma) &= \int_\Theta \{ \log(1 - t^* \Gamma(\nu)) + t^* \log(1 + \Gamma(\nu)) \} d\nu \\ &\quad + \tau(1 + t^*), \end{aligned}$$

$$\begin{aligned} \kappa_m(\Gamma) &= \int_\Theta \{ \log(1 - t^* \Gamma(\nu)) + t^* \log(1 + \Gamma(\nu)) \} d\nu \\ &\quad + \tau t^*, \end{aligned}$$

where t^* is the unique solution to

$$\tau + \int_\Theta \log(1 + \Gamma(\nu)) d\nu = \int_\Theta \frac{\Gamma(\nu)}{1 - t^* \Gamma(\nu)} d\nu. \quad (21)$$

A. Centralized Error Exponents

Theorem 3 (CD Error Exponents). Consider the LLR test in (3) with a fixed threshold τ , the spectral density $\Gamma_{\text{CD}} \triangleq \Gamma_{\text{CD}}(\nu) = \frac{\phi(\nu)}{\sigma_v^2}$, $\Theta = [0, 1]$ and assume that $\phi(\nu)$ is in the Wiener class. Then, the false alarm and miss error exponents for the hypothesis testing problem in (1) are:

$$\kappa_{fa}^{\text{CD}} = \begin{cases} \kappa_{fa}(\Gamma_{\text{CD}}) & \text{if } \tau > m_0(\Gamma_{\text{CD}}), \\ 0 & \text{if } \tau \leq m_0(\Gamma_{\text{CD}}), \end{cases} \quad (22)$$

$$\kappa_m^{\text{CD}} = \begin{cases} \kappa_m(\Gamma_{\text{CD}}), & \text{if } \tau < m_1(\Gamma_{\text{CD}}), \\ 0 & \text{if } \tau \geq m_1(\Gamma_{\text{CD}}), \end{cases} \quad (23)$$

Proof: See App. I. ■

Corollary 1. Consider the LLR test in (3). The miss error exponent subject to $P_{fa}^n \leq \alpha$ with $\alpha \in (0, 1)$ is

$$\kappa_{m,\alpha}^{\text{CD}} = \int_0^1 \left\{ \frac{1}{1 + \Gamma_{\text{CD}}(\nu)} + \log(1 + \Gamma_{\text{CD}}(\nu)) - 1 \right\} d\nu. \quad (24)$$

Proof: Evaluate Th. 3 with $\tau = m_0(\Gamma_{\text{CD}}) + \epsilon$, where $\epsilon > 0$ is arbitrary small. See [33, Prop. 2] for a detailed proof. ■

B. Decentralized Error Exponents

Consider the setup in (4), and the LLR test in (5). Bearing in mind the asymptotic behavior of the covariance matrices of \mathbf{z} under \mathcal{H}_0 and \mathcal{H}_1 , we formulate the following lemma:

Lemma 2 (Decentralized Hypothesis Testing). Under either PCS-MAC or PFS-MAC, the decision at the FC asymptotically consists on choosing one of the following PSDs, with $\{\gamma_{k'}^n\}$ as in (9), and $\nu \in \Theta_\beta$:

$$\begin{cases} \mathcal{H}_1: & h(\nu) = (\phi(\nu) + \sigma_v^2)\xi(\nu) + \sigma_w^2 \\ \mathcal{H}_0: & h(\nu) = \sigma_v^2\xi(\nu) + \sigma_w^2. \end{cases} \quad (25)$$

Proof: See App. II-A. ■

The following theorem establishes the error exponents for the decentralized strategies presented in Section III. Both strategies PCS-MAC and PFS-MAC allow to improve the performance of detection by careful selection of $\xi(\nu)$.

Theorem 4 (DD Error Exponents). Consider the hypothesis testing problem in (25), where $\xi(\nu)$ is a fixed asymptotic energy profile that satisfies (10). Consider also that the asymptotic fraction of DoF is limited to β . Assume that $\phi(\nu)$ is in the Wiener class and that the threshold of the test is fixed to τ . Let $\Gamma_{DD} \triangleq \Gamma_{DD}(\nu) = \frac{\xi(\nu)\phi(\nu)}{\xi(\nu)\sigma_v^2 + \sigma_w^2}$. Then, the strategies PCS-MAC and PFS-MAC have the same false alarm and miss error exponents, and they are:

$$\kappa_{fa}^{DD} = \begin{cases} \kappa_{fa}(\Gamma_{DD}) & \text{if } \tau > m_0(\Gamma_{DD}), \\ 0 & \text{if } \tau \leq m_0(\Gamma_{DD}), \end{cases} \quad (26)$$

$$\kappa_m^{DD} = \begin{cases} \kappa_m(\Gamma_{DD}), & \text{if } \tau < m_1(\Gamma_{DD}), \\ 0 & \text{if } \tau \geq m_1(\Gamma_{DD}), \end{cases} \quad (27)$$

where t^* is the unique solution to (21) and $m_0(\cdot)$ and $m_1(\cdot)$ are given by (19) and (20) with $\Gamma = \Gamma_{DD}$, $\Theta = \Theta_\beta$.

Proof: See App. II-B. ■

Corollary 2 (DD Miss Error Exponent and Optimum Energy Profile). Consider the hypothesis testing problem in (25). Assume that $\phi(\nu)$ is in the Wiener class, the fraction of DoF is limited to β , and the energy constraint is given by (10). The optimal miss error exponent subject to $P_{fa}^n \leq \alpha$, with $\alpha \in (0, 1)$ fixed, considering orthogonal schemes is achieved with either the PCS-MAC or PFS-MAC scheme and is given by

$$\kappa_{m,\alpha}^{DD} = \int_{\Theta_{\beta*}} \left\{ \frac{1}{1 + \Gamma_{DD}(\nu)} + \log(1 + \Gamma_{DD}(\nu)) - 1 \right\} d\nu, \quad (28)$$

where $\Theta_{\beta*} \subseteq \Theta_\beta$ is the support of the optimal energy profile $\xi_{OEP}(\nu) = \max\{\hat{\xi}_1(\nu), \hat{\xi}_2(\nu), \hat{\xi}_3(\nu), 0\}$ and $\hat{\xi}_i(\nu)$, $i = 1, 2, 3$, are the roots of the following cubic equation:

$$a_3(\nu)\hat{\xi}^3 + a_2(\nu)\hat{\xi}^2 + a_1(\nu)\hat{\xi} + a_0 = 0 \quad (29)$$

with coefficients

$$\begin{aligned} a_0 &= \lambda^* \sigma_w^6 \\ a_1(\nu) &= \sigma_w^2(-\phi(\nu)^2 + \lambda^* \sigma_w^2(2\phi(\nu) + 3\sigma_v^2)) \\ a_2(\nu) &= \lambda^* \sigma_w^2(\phi(\nu)^2 + 4\phi(\nu)\sigma_v^2 + 3\sigma_v^4) \\ a_3(\nu) &= \lambda^* \sigma_v^2(\phi(\nu) + \sigma_v^2)^2, \end{aligned} \quad (30)$$

where λ^* is the Lagrange multiplier that satisfies the energy constraint (10) with equality. The closed-form solution of $\xi_{OEP}(\nu)$ is shown in (31), at the top of the next page.

Proof: See App. II-C. ■

C. Suboptimal Energy Profiles

In this section we define three energy profiles: i) constant energy profile, $\xi_{CEP}(\nu) \triangleq \frac{E_t}{(\sigma_v^2 + \sigma_s^2)}$, where all the sensors transmit using the same gain for all channel uses; ii) spectral energy profile, $\xi_{SEP}(\nu) \triangleq \frac{E_t \phi(\nu)}{\sigma_s^2(\sigma_v^2 + \sigma_s^2)}$, that reproduces the shape of $\phi(\nu)$ and therefore, allocates more energy to the frequencies where the process under \mathcal{H}_1 concentrates more power; iii) ON/OFF energy profile, $\xi_{ON/OFF-EP}(\nu) \triangleq \frac{E_t}{(\sigma_v^2 + \sigma_s^2)\beta^*} \mathbb{1}(\nu \in \Theta_{\beta*})$, with $\beta^* = L(\Theta_{\beta*}) \leq \beta$, where each sensor transmits with constant gain if $\nu \in \Theta_{\beta*}$ and stays silent otherwise. In i) and ii) we assume that the sensors do not have access to the transmitted modes set and no DoF compression is possible ($\beta = 1$). In iii) and in the optimal energy profile, the sensors know the transmitted modes set.

VI. NUMERICAL RESULTS FOR 1D

In this section, we evaluate the miss error exponent for two complex Gaussian correlated auto-regressive moving average (ARMA) processes with PSD given by $\phi(\nu) = \sigma_{in}^2 \left| \frac{\sum_{k=0}^M b_k e^{-j2\pi\nu k}}{\sum_{k=0}^N a_k e^{-j2\pi\nu k}} \right|^2$, where M and N are the degrees of the numerator and denominator polynomials, respectively, and σ_{in}^2 is selected such that the variance of the process is σ_s^2 . The simulated processes (with even PSD) are plotted in Fig. 3 with coefficients shown in Table I. We will referred to them as PSD1 and PSD2. Let $\text{SNR}_M = \sigma_s^2/\sigma_v^2$ and $\text{SNR}_C = E_t/\sigma_w^2$

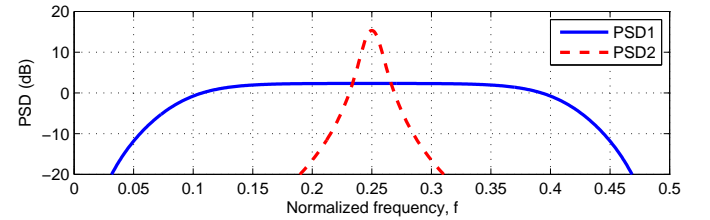


Fig. 3. PSDs (in dB) with variance $\sigma_s^2 = 1$. Only positive normalized frequencies are plotted.

be the measurement and communication signal-to-noise ratios, respectively. We use $\text{SNR}_M = 5$ dB, $\sigma_s^2 = 1$ and $E_t = 1$ on all figures of this section. In Fig. 4, we show the behavior of the miss error exponent for processes PSD1 and PSD2 against SNR_C when $P_{fa}^n \leq \alpha$. In both cases the decentralized detectors (DD) approach the centralized detector (CD) performance when SNR_C is high enough. We see that the optimum scheme DD-OEP allows to save a significant amount of energy when the communication signal-to-noise ratio is low. This is the desirable operation regime of a wireless sensor network with massive amount of nodes since it would extend the useful life of the WSN, avoid maintenance action for changing the battery of the nodes, or even elude network reconfiguration when nodes run out of energy.

$$\begin{aligned}
\xi_{\text{OEP}}(\nu) = & \{ (1/(62^{1/3} \lambda (\phi^2 v + 2\phi v^2 + v^3))) ((1 - j\sqrt{3})(2\lambda^3 \phi^6 w^3 + 6\lambda^3 \phi^5 v w^3 + 6\lambda^3 \phi^4 v^2 w^3 + 2\lambda^3 \phi^3 v^3 w^3 + 9\lambda^2 \phi^6 v w^2 \\
& + 54\lambda^2 \phi^5 v^2 w^2 + 108\lambda^2 \phi^4 v^3 w^2 + 90\lambda^2 \phi^3 v^4 w^2 + 27\lambda^2 \phi^2 v^5 w^2 + (4(3\lambda w(\phi^2 v + 2\phi v^2 + v^3)(2\lambda \phi w + 3\lambda v w - \phi^2) \\
& - \lambda^2 w^2 (\phi^2 + 4\phi v + 3v^2)^2)^3 + (2\lambda^3 \phi^6 w^3 + 6\lambda^3 \phi^5 v w^3 + 6\lambda^3 \phi^4 v^2 w^3 + 2\lambda^3 \phi^3 v^3 w^3 + 9\lambda^2 \phi^6 v w^2 + 54\lambda^2 \phi^5 v^2 w^2 \\
& + 108\lambda^2 \phi^4 v^3 w^2 + 90\lambda^2 \phi^3 v^4 w^2 + 27\lambda^2 \phi^2 v^5 w^2)^2)^{2/3} - ((1 + j\sqrt{3})(3\lambda w(\phi^2 v + 2\phi v^2 + v^3)(2\lambda \phi w + 3\lambda v w - \phi^2) \\
& - \lambda^2 w^2 (\phi^2 + 4\phi v + 3v^2)^2)) / (32^{2/3} \lambda (\phi^2 v + 2\phi v^2 + v^3) (2\lambda^3 \phi^6 w^3 + 6\lambda^3 \phi^5 v w^3 + 6\lambda^3 \phi^4 v^2 w^3 + 2\lambda^3 \phi^3 v^3 w^3 \\
& + 9\lambda^2 \phi^6 v w^2 + 54\lambda^2 \phi^5 v^2 w^2 + 108\lambda^2 \phi^4 v^3 w^2 + 90\lambda^2 \phi^3 v^4 w^2 + 27\lambda^2 \phi^2 v^5 w^2 + (4(3\lambda w(\phi^2 v + 2\phi v^2 + v^3) \\
& \times (2\lambda \phi w + 3\lambda v w - \phi^2) - \lambda^2 w^2 (\phi^2 + 4\phi v + 3v^2)^2)^3 + (2\lambda^3 \phi^6 w^3 + 6\lambda^3 \phi^5 v w^3 + 6\lambda^3 \phi^4 v^2 w^3 + 2\lambda^3 \phi^3 v^3 w^3 \\
& + 9\lambda^2 \phi^6 v w^2 + 54\lambda^2 \phi^5 v^2 w^2 + 108\lambda^2 \phi^4 v^3 w^2 + 90\lambda^2 \phi^3 v^4 w^2 + 27\lambda^2 \phi^2 v^5 w^2)^2)^{2/3} - (w(\phi^2 + 4\phi v + 3v^2)) \\
& / (3(\phi^2 v + 2\phi v^2 + v^3))) \} \mathbb{1}(D(\nu) \geq 0).
\end{aligned}
\tag{31}$$

$$D(\nu) = -(\lambda \phi^2 + 2\lambda \phi v + \lambda v^2)(4\lambda^2 \phi^5 w^5 - \lambda \phi^6 w^4 + 18\lambda \phi^5 v w^4 + 27\lambda \phi^4 v^2 w^4 - 4\phi^6 v w^3). \tag{32}$$

Fig. 2. Optimal energy profile and the discriminant of (29), $D(\nu)$. For readability, we define $\phi \triangleq \phi(\nu)$, $w \triangleq \sigma_w^2$, $v \triangleq \sigma_v^2$ and $\lambda \triangleq \lambda^*$.

Coefficients PSD1: $\sigma_m^2 = 1.70$				
b_0	b_1	b_2	b_3	b_4
.39	0	-.78	0	.39
a_0	a_1	a_2	a_3	a_4
1	0	-.37	0	.19

Coefficients PSD2: $\sigma_m^2 = 2.37 \cdot 10^{-5}$				
b_0	b_1	b_2	b_3	b_4
3	0	-6	0	3
a_0	a_1	a_2	a_3	a_4
1	0	1.82	0	0.83

TABLE I

ARMA COEFFICIENTS TO GENERATE BOTH PSDS WITH VARIANCE $\sigma_s^2 = 1$ AND $N = M = 4$.

Table II shows the energy savings of the optimum scheme DD-OEP and the suboptimal energy profile DD-SEP with respect to (wrt) the constant energy profile DD-CEP for several miss error exponents, and for both PSDs.

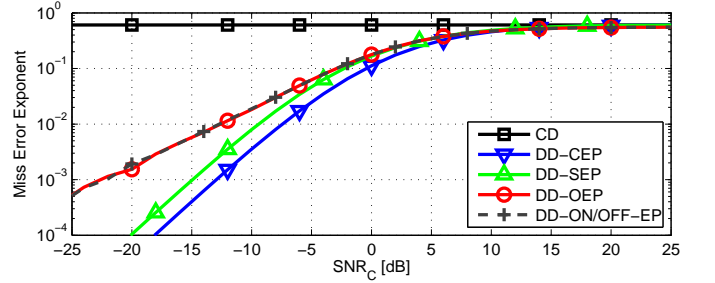
A remarkable result is obtained when the scheme DD-ON/OFF-EP is used. Comparing with the optimal strategy, DD-OEP, we observe that the differences between the exponents obtained with both strategies are negligible. This shows

Energy gap (in dB) wrt CEP for PSD1		
κ_m	SEP	OEP, $\beta = 0.6$
10^{-1}	2	2.5
10^{-2}	2	5
10^{-3}	2	10

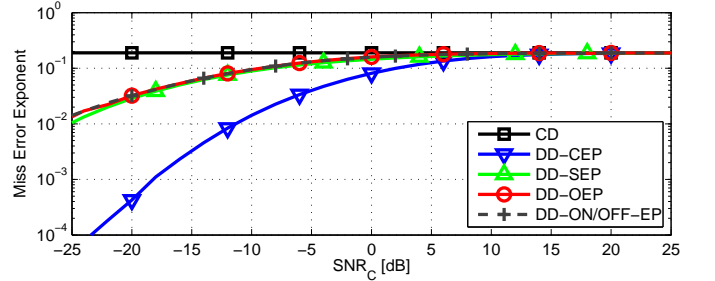
Energy gap (in dB) wrt CEP for PSD2		
κ_m	SEP	OEP, $\beta = 0.2$
10^{-1}	10.5	11
10^{-2}	14	15
10^{-3}	>14	>15

TABLE II

ENERGY SAVING ALLOCATING ENERGY IN THE SENSORS.



(a) PSD1, $\beta = 0.6$ for OEP and ON/OFF.



(b) PSD2, $\beta = 0.2$ for OEP and ON/OFF.

Fig. 4. Error exponents for CD and DD detectors against SNR_C (dB) for process PSD1 and PSD2.

that knowing Θ_{β^*} is much more important than knowing the optimal gains $\{\gamma_{k'}^n\}$. This observation leads to a very simple strategy: at the beginning of the detection process, the sensors are communicated the optimal set Θ_{β^*} (e.g., the FC broadcasts it through a low-rate feedback channel); then, on all channel uses, each node transmits with the same gain using PFS-MAC.

In Fig. 5, we plot the optimum fraction of DoF β^* constrained to the allowed fraction of DoF β against this constraint for processes PSD1 and PSD2. For low values of β , β^* increases linearly with slope one ($\beta^* \leq \beta$) up to a certain value where β^* saturates. The saturation effect is observed for different levels depending on SNR_C and on the frequency

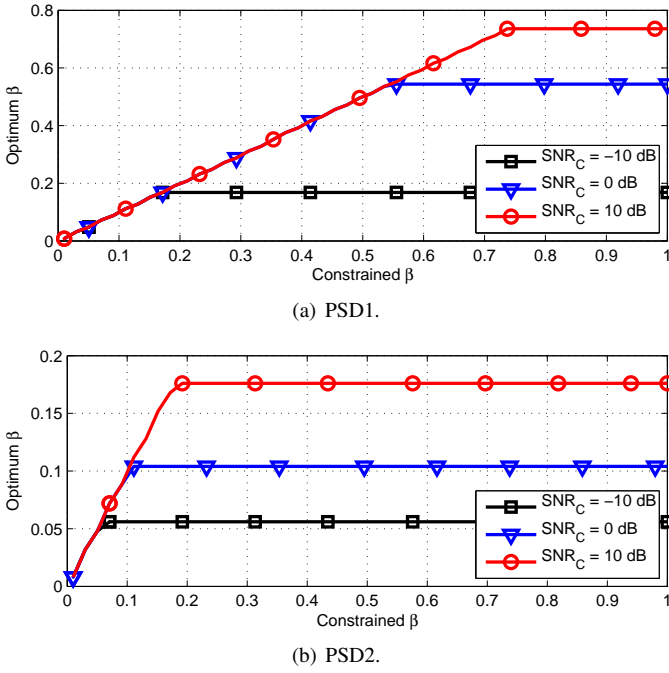


Fig. 5. Optimum constrained fraction of DoF against the constrained fraction of DoF.

selectivity of the process, which is related to its correlation. Thus, a strongly correlated process (PSD2) needs relatively less channel uses than a weakly correlated process (PSD1). When SNR_C is high, saturation occurs for high values of β . Conversely, if SNR_C is low, the optimum scheme uses the available energy E_t in each sensor to transmit more reliably a reduced set of frequencies where the PSD of the process is high. As a remark, the asymptotically optimal orthogonal strategy PFS-MAC not only allows to save a valuable amount of energy in the sensors but also allows to save channel uses (i.e., bandwidth, detection delay). In Fig. 6, we plot the optimum energy profiles as a function of the normalized frequency for both processes and for several values of SNR_C, when the constraint on the fraction of DoF is inactive ($\beta = 1$). Note that these figures are closely related to Fig. 5 since the Lebesgue measure of the support of the optimum energy profile is indeed β^* .

VII. 2 AND 3 DIMENSIONAL WSNs

In this section we enlarge the model to networks with up to 2 dimensions in space and we include the time dimension. The extension to more dimensions is straightforward. The measurement taken by the sensor at coordinate (i, j) at time m under each hypothesis is:

$$\begin{cases} \mathcal{H}_1: & x_{ij}(m) = s_{ij}(m) + v_{ij}(m), & i = 1, \dots, n_1, \\ \mathcal{H}_0: & x_{ij}(m) = v_{ij}(m), & j = 1, \dots, n_2, \\ & & m = 1, \dots, n_3. \end{cases}$$

where the parameters n_1 , n_2 , and n_3 are chosen for describing one, two, or three dimensional networks. In these networks, sensors are distributed along a line (1D) or in space (2D), and they take several measurements during different time

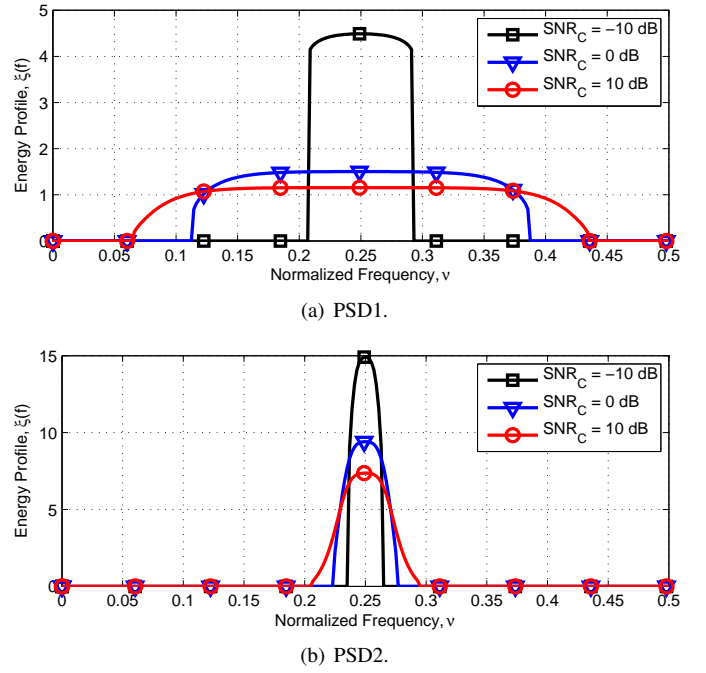


Fig. 6. Energy profiles for DD-OEP detector against the normalized frequency for process with PSD1 and PSD2 without DoF constraint ($\beta = 1$). Only $[0, 0.5]$ frequency interval is shown due to the even symmetry of $\xi(\nu)$ inherited from $\phi(\nu)$.

instants (time dimension). We assume that $s_{ij}(m)$ is a zero-mean circularly-symmetric complex Gaussian stationary process with variance σ_s^2 and power spectral density (PSD) $\phi(\nu)$, where $\nu \in [0, 1]^p$ is the vector of frequencies normalized to the interval $[0, 1]$ and p is the dimension of the network. $v_{ij}(m)$ is zero-mean circularly-symmetric complex white Gaussian noise independent of $s_{ij}(m)$ with variance σ_v^2 . Therefore, $x_{ij}(m)$ is Gaussian distributed under each hypothesis.

All sensors transmit synchronously over a MAC, and the received signal at the FC during the k' -th channel use is:

$$z_{k'}^n = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} g_{ijk'}(\{x_{ij}(m)\}_{m=1}^{n_3}) + w_{k'}, \quad k' = 1, \dots, n', \quad (33)$$

where $g_{ijk'}(\{x_{ij}(m)\}_{m=1}^{n_3})$ is the symbol transmitted by the sensor located at (i, j) during channel use k' using the encoding function $g_{ijk'}(\cdot)$, $w_{k'}$ is a zero-mean circularly-symmetric complex white Gaussian noise independent of everything else with variance σ_w^2 , and n' is the number of MAC uses. As in the one dimensional case, we will consider linear encoding functions only. To each 3-tuple index (i, j, m) , we associate the following unwound index

$$k = (i - 1)n_2n_3 + (j - 1)n_3 + m \quad (34)$$

that takes values $k = 1, \dots, n$ with $n = n_1n_2n_3$. This is a one to one mapping, i.e., each 3-tuple index (i, j, m) can be recovered from the index k and vice versa. We denote this as $k \leftrightarrow (i, j, m)$. If we define the following column vectors: $\mathbf{s} = [s_1^n, \dots, s_n^n]^T$, $\mathbf{v} = [v_1^n, \dots, v_n^n]^T$ and $\mathbf{x} = [x_1^n, \dots, x_n^n]^T$, the vector of measurements at the FC are expressed as in (4)

and hence, its statistic results (5). However, we assume now that the covariance matrix Σ_n is a p -level Toeplitz matrix [34, Sec. 6.4] instead of a regular Toeplitz matrix. The precoding matrix for the p -dimensional (p -D) PCS-MAC strategy has the same expression (6) although the eigenvectors of the p -level Toeplitz matrix Σ_n change. However, we need to redefine the precoding matrix for the p -D PFS-MAC strategy.

Definition 9 (p -D PFS-MAC). For each $n = n_1 \times \dots \times n_p$, denote by (j_1, j_2, \dots, j_n) a permutation of $\{1, 2, \dots, n\}$ such that $\phi[j_1] \geq \phi[j_2] \geq \dots \geq \phi[j_n]$ where $j_{k'} \leftrightarrow (i'_1, \dots, i'_p)$, $i'_l \in [1, n_l]$, $l = 1, \dots, p$ and $\phi[j_{k'}]$ is defined as a sample of the p -D PSD: $\phi[j_{k'}] \triangleq \phi\left(\frac{i'_1-1}{n_1}, \dots, \frac{i'_p-1}{n_p}\right)$ with $k' = 1, \dots, n'$. The precoding matrix of the p -D PFS-MAC strategy is

$$C_{nn'} = F_{nn'} \Delta_{n'}, \quad (35)$$

where $F_{nn'} = [f_{j_1}^n, \dots, f_{j_{n'}}^n]$ is a sub-matrix of the p -D DFT matrix of size $n \times n$, i.e., $f_{k'}^n = [f_{1k'}^n, f_{2k'}^n, \dots, f_{nk'}^n]^T$ with $f_{kk'}^n = \frac{1}{\sqrt{n}} \exp(j2\pi \sum_{l=1}^p \frac{(i_l-1)(i'_{l,k'}-1)}{n_l})$, $k = 1, \dots, n$; $k \leftrightarrow (i_1, \dots, i_p)$, and $\Delta_{n'} = \text{diag}(\gamma_1^n, \dots, \gamma_{n'}^n)$.

Similar to the 1D case, the objective of both p -D PCS-MAC and p -D PFS-MAC is to communicate the most important modes (frequencies) of the random process to the FC. Under both strategies, each sensor needs to know the complete vector of time measurements to transmit during each channel use a different p -tuple index $(i'_{1,k'}, \dots, i'_{p,k'})$. Once the FC has the n' measurements, it builds the statistic (5) to make a decision.

We need now a multi-dimensional version of the Toeplitz theorem.

Theorem 5 (Toeplitz Distribution for p -D processes). For a Hermitian p -level Toeplitz matrix Σ_n generated by the spectral density $\phi(\boldsymbol{\nu})$ which belongs to the Wiener class, with $\boldsymbol{\nu} = (\nu_1, \dots, \nu_p)$ and multilevel index $\mathbf{n} = (n_1, \dots, n_p)$, $n = n_1 \times \dots \times n_p$, let $\{\lambda_k^n\}_{k=1}^n$ be the eigenvalues of Σ_n contained on the interval $[\delta_1, \delta_2]$, let $F(\cdot)$ be a continuous function defined on $[\delta_1, \delta_2]$ and assume that $\int_{\boldsymbol{\nu}: \phi(\boldsymbol{\nu})=\delta} F(\phi(\boldsymbol{\nu})) d\boldsymbol{\nu} = 0$ for any $\delta \in [\delta_1, \delta_2] \cap \Delta$ where $\Delta \subseteq [0, \infty)$ then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n F(\lambda_k^n \in \Delta) = \int_{\phi^{-1}(\Delta)} F(\phi(\boldsymbol{\nu})) d\boldsymbol{\nu}.$$

where $n \rightarrow \infty$ means that all components of \mathbf{n} tend to infinity simultaneously.

Proof: See [34, Th. 6.4.1] and [30, Corollary 4.1].

Using Th. 5 and Th. 2 together with the p -D PCS-MAC and p -D PFS-MAC strategies we obtain the same results given in Section V. This is summarized in the following theorem.

Theorem 6 (DD Error Exponents for p -D networks). The error exponents for the p -D PCS-MAC and p -D PFS-MAC strategies are given by Th. 4 and Corollary 2 under the same hypotheses considering now that the normalized frequency is a p -dimensional variable $\boldsymbol{\nu}$ and Θ is a p -dimensional set.

Proof: Apply Th. 4 using (34). ■

The same energy profiles of subsection V-C are defined if we consider again that the normalized frequency is a p -dimensional variable $\boldsymbol{\nu}$ and Θ is a p -dimensional set.

VIII. NUMERICAL EXPERIMENT FOR 2D

In this section we compare the theoretical results with a Monte Carlo simulation for detecting a 2D correlated process described by the following partial differential equation:

$$\tilde{a}_x \frac{\partial^2 s(x, y)}{\partial x^2} + \tilde{a}_y \frac{\partial^2 s(x, y)}{\partial y^2} + \tilde{a}_0 s(x, y) = q(x, y) \quad (36)$$

where $q(x, y)$ is the random source, assumed to be white and Gaussian with PSD σ_q^2 , and \tilde{a}_x , \tilde{a}_y and \tilde{a}_0 are constants. Using a second order approximation, we discretize (36),

$$\frac{\partial^2 s(x, y)}{\partial x^2} \approx \frac{s(x+h, y) - 2s(x, y) + s(x-h, y)}{h^2} \quad (37)$$

$$\frac{\partial^2 s(x, y)}{\partial y^2} \approx \frac{s(x, y+h) - 2s(x, y) + s(x, y-h)}{h^2} \quad (38)$$

where h is the discretization step in both directions. The discrete equation results

$$a_x(s_{i+1,j} + s_{i-1,j}) + a_y(s_{i,j+1} + s_{i,j-1}) + a_0 s_{i,j} = q_{i,j}, \\ i = 1, \dots, n_1, j = 1, \dots, n_2.$$

with $a_x = \tilde{a}_x/h^2$, $a_y = \tilde{a}_y/h^2$ and $a_0 = \tilde{a}_0 - 2(\tilde{a}_x + \tilde{a}_y)/h^2$. The PSD of the process is

$$\phi(\omega_x, \omega_y) = \frac{\sigma_q^2}{(a_0 + 2a_x \cos(\omega_x) + 2a_y \cos(\omega_y))^2} \quad (39)$$

and its variance is σ_s^2 . If we organize the samples of the process in a vector as in Section VII, (39) results a linear system: $A_n \mathbf{s}_n = \mathbf{q}_n$. In there, \mathbf{s}_n and \mathbf{q}_n are vectors of dimension $n = n_1 n_2$, and A_n is a two-level banded Toeplitz matrix of size $n \times n$ with the sub-block matrix B_{n_1} :

$$A_n = \begin{bmatrix} B_{n_1} & a_x I_{n_1} & & \\ a_x I_{n_1} & B_{n_1} & \ddots & \\ & \ddots & \ddots & \ddots \end{bmatrix}, B_{n_1} = \begin{bmatrix} a_0 & a_y & & \\ a_y & a_0 & \ddots & \\ & \ddots & \ddots & \ddots \end{bmatrix},$$

where only non-zero elements are indicated. Assuming that A_n is non-singular, the covariance matrix of the process is

$$\Sigma_n = \sigma_q^2 (A_n A_n^T)^{-1} \quad (40)$$

which is in general a non-Toeplitz matrix. However, A_n is asymptotically a circulant matrix. The product of two circulant matrices, as well as the inverse of a circulant matrix, are circulant matrices [30, p. 50, 63 and 67]. Circulant matrices are a particular case of Toeplitz matrices. Therefore, Σ_n is asymptotically a 2-level Toeplitz matrix and we apply Th. 6 to compute the miss error exponent subject to any fixed level of false alarm error probability $\alpha \in (0, 1)$.

In Fig. 7 we show the estimation of the miss error exponent as a function of the number of sensors in the network using the Monte Carlo method with the following parameters: $\text{SNR}_C = -10$ dB, $a_0 = -5$, $a_x = a_y = 1$, 10^6 experiments and $\alpha = 10^{-2}$. We consider that the sensors are placed in a regular

square grid, i.e., $n_1 = n_2$ and the total amount of sensors in the network is $n = n_1 n_2$. The threshold of the test is modified for each n in order to keep the false alarm error probability constant. We also plot the theoretical miss error exponent. We observe that both estimated miss error exponents converge to their corresponding theoretical values. We also note that both decentralized schemes PCS-MAC and PFS-MAC have almost the same performance for the number of sensor considered in the figure, which validates the circulant approximation of the product of Toeplitz matrices in (40).

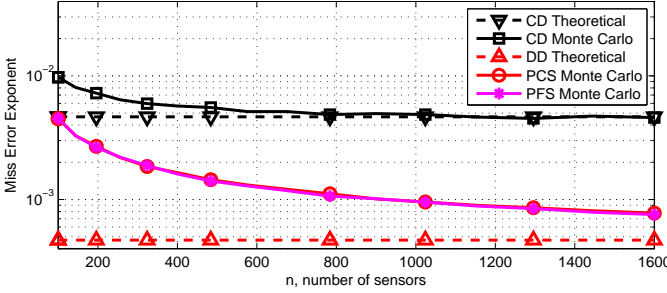


Fig. 7. Miss error probability vs. the amount of sensors in the network.

IX. CONCLUSIONS

We have proposed several schemes for distributed detection of circularly-symmetric complex Gaussian random processes with arbitrary correlation function both in space and time. These schemes take into account possible correlated measurements and use this correlation beneficially at the FC to build an appropriate statistic to make a decision. Considering a multiple access channel and imposing bandwidth and per-sensor energy constraints, we have obtained the optimal orthogonal scheme in terms of the miss error exponent. We have also shown that one of the proposed schemes is particularly attractive for WSNs with low-cost and energy-limited nodes because it requires only the set of modes to be transmitted, obtains significant energy saving in the low signal-to-noise ratio regime, and achieves a close-to-optimal (with negligible loss) performance in terms of miss error exponent.

APPENDIX I

PROOF OF THEOREM 3: CD ERROR EXPONENTS

The asymptotic mean of the centralized LLR statistic defined in (3) under the hypothesis \mathcal{H}_i , $i = 0, 1$ is

$$\begin{aligned} m_i^{\text{CD}} &= \lim_{n \rightarrow \infty} \mathbb{E}_i(T_c^n) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \left(\text{tr}(\Sigma_{i,n}(\Sigma_{0,n}^{-1} - \Sigma_{1,n}^{-1})) - \log \frac{\det \Sigma_{1,n}}{\det \Sigma_{0,n}} \right). \end{aligned}$$

Considering that the eigenvalues of $\Sigma_{0,n}$ and $\Sigma_{1,n}$ are $\{\sigma_v^2\}$ and $\{\lambda_k^n + \sigma_v^2\}$, respectively, and defining $\rho_{\text{CD},k}^n = \frac{\lambda_k^n}{\sigma_v^2}$, we obtain $m_0^{\text{CD}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n 1 - \frac{1}{1 + \rho_{\text{CD},k}^n} - \log(1 + \rho_{\text{CD},k}^n)$. Using Th. 1 with $\Delta = [\delta_1, \delta_2]$, which implies $\phi^{-1}(\Delta) = [0, 1]$, we finally have $m_0^{\text{CD}} = m_0(\Gamma_{\text{CD}})$ defined (19).

To compute the error exponents we first need to verify the assumptions of Th. 2 with $h_0 \triangleq h_0(\nu) = \sigma_v^2 \forall \nu$ and $h_1 \triangleq h_1(\nu) = \phi(\nu) + \sigma_v^2$:

- $A_1)$ h_0 is a positive constant and $\log h_0$ belongs to $L^1([0, 1])$ trivially. $\log h_1 \in L^1([0, 1])$ is easily proved by noting that $\text{ess inf}(\phi(\nu) + \sigma_v^2) > 0$ and $\text{ess sup}(\phi(\nu) + \sigma_v^2) < \infty$ because $\phi(\nu)$ is a power spectral density in the Wiener class, and therefore, it is essentially bounded.
- $A_2)$ $h_0/h_1 = 1/(1 + \Gamma_{\text{CD}}) \in L^\infty([0, 1])$ given that $\text{ess inf} \Gamma_{\text{CD}} \geq 0$ because Γ_{CD} is a spectral density. $h_1/h_0 = 1 + \Gamma_{\text{CD}} \in L^\infty([0, 1])$ is easily proved by considering again that $\phi(\nu)$ belongs to the Wiener class.

Given that (A_1) and (A_2) are satisfied, the error exponents are obtained through the Fenchel-Legendre transform $\Lambda_{\text{CD},i}^*(x)$ of the LMGF $\Lambda_{\text{CD},i}(t) = \log \mathbb{E}_i[e^{tT_c^n}]$, $i = 0, 1$. The following properties can be verified:

Properties 1 (LMGF and its Fenchel Legendre Transform).

- $P_1)$ $\Lambda_{\text{CD},1}(-1) = \Lambda_{\text{CD},1}(0) = 0$.
- $P_2)$ $\Lambda'_{\text{CD},1}(-1) = m_0^{\text{CD}}$ and $\Lambda'_{\text{CD},1}(0) = m_1^{\text{CD}}$.
- $P_3)$ $\Lambda_{\text{CD},1}(t)$ is a convex function and then $\Lambda_{\text{CD},1}(t) = \Lambda_{\text{CD},1}((1+t)0 - t(-1)) \leq (1+t)\Lambda_{\text{CD},1}(0) - t\Lambda_{\text{CD},1}(-1) = 0 \forall t \in [-1, 0]$.
- $P_4)$ As the Fenchel-Legendre transform given in (18) transforms convex functions into convex functions, $\Lambda_{\text{CD},1}^*(x)$ is a convex function. Moreover, $\Lambda_{\text{CD},1}^*(x)$ has a minimum of value 0 at $x = m_1^{\text{CD}}$ meaning that $\Lambda_{\text{CD},1}^*(x)$ is decreasing if $x < m_1^{\text{CD}}$.
- $P_5)$ $\Lambda_{\text{CD},0}(t) = \Lambda_{\text{CD},1}(t-1)$ and $\Lambda_{\text{CD},0}^*(x) = \Lambda_{\text{CD},1}^*(x) + x$. Therefore, $\Lambda_{\text{CD},0}^*(x)$ is also a convex function with a minimum of value 0 at $x = m_0^{\text{CD}}$ and it is an increasing function if $x > m_0^{\text{CD}}$.

Using properties (P_1) - (P_5) and (16) when $G = [\tau, \infty)$, $k_{fa}^{\text{CD}} = \Lambda_{\text{CD},0}^*(\tau)$ if $\tau > m_0^{\text{CD}}$ and $k_{fa}^{\text{CD}} = 0$ if $\tau \geq m_0^{\text{CD}}$. Moreover, when $G = (-\infty, \tau]$, $k_m^{\text{CD}} = \Lambda_{\text{CD},1}^*(\tau)$ if $\tau < m_1^{\text{CD}}$ and $k_m^{\text{CD}} = 0$ if $\tau \geq m_1^{\text{CD}}$. The usual interval of interest for the threshold τ is $[m_0^{\text{CD}}, m_1^{\text{CD}}]$ and because (P_1) and (P_2) hold, the interval of optimization of t in (18) can be restricted to $[-1, 0]$. Due to the convexity of $\Lambda_{\text{CD},1}(t)$ there exists a unique t^* that solves (18) and satisfies (21), which is obtained from deriving $\tau t - \Lambda_{\text{CD},1}(t)$. Then, the error exponents are (22) and (23).

APPENDIX II

DECENTRALIZED ERROR EXPONENTS

A. Proof of Lemma 2: Decentralized Hypothesis Testing

We first consider the PCS-MAC scheme given in Def. 2. The covariance matrices under \mathcal{H}_0 and \mathcal{H}_1 are, respectively,

$$\begin{aligned} \Xi_{0,n'} &= \sigma_v^2 \Delta_{n'}^2 + \sigma_w^2 I_{n'}, \\ \Xi_{1,n'} &= (\text{diag}(\lambda_1^n, \dots, \lambda_{n'}^n) + \sigma_v^2 I_{n'}) \Delta_{n'}^2 + \sigma_w^2 I_{n'}. \end{aligned} \quad (41)$$

Let $\theta_{i,k}$ be the k -th element of the diagonal matrix $\Xi_{i,n'}$, $i = 0, 1$. It is easy to prove that $\text{diag}(\theta_{0,1}, \dots, \theta_{0,n'}) \sim \text{diag}(h_i(\frac{0}{n}), \dots, h_i(\frac{n'-1}{n}))$. To recover the original dimension of the problem, define the n -dimensional vector \tilde{z} as

the zero padding of the n' -dimensional vector of measurements \mathbf{z} , i.e., $\tilde{\mathbf{z}} = [\mathbf{z}^T \mathbf{0}^T]^T$. The covariance matrix of $\tilde{\mathbf{z}}$, $\text{diag}(\theta_{i,1}, \dots, \theta_{i,n'}, 0, \dots, 0)$, is asymptotically equivalent to $\text{diag}(h_i(\frac{0}{n}), \dots, h_i(\frac{n'-1}{n}), 0, \dots, 0)$. Consider the following transformation $\mathbf{y} = F_n \tilde{\mathbf{z}}$, where F_n is the DFT matrix of size $n \times n$. It is well known that applying an invertible transformation (in particular, the orthogonal DFT matrix) to the data does not modify the performance of the statistic. Because matrix multiplication preserves asymptotic equivalence of matrices [30, Th. 2.1 (3)], we have that the covariance matrix of \mathbf{y} is asymptotically circulant, i.e.,

$$F_n \text{diag}(\theta_{i,1}, \dots, \theta_{i,n'}, 0, \dots, 0) F_n^H \sim F_n \text{diag}(h_i(0/n), \dots, h_i((n'-1)/n), 0, \dots, 0) F_n^H = B_n(\tilde{h}_i),$$

where $B_n(\tilde{h}_i)$ is the circulant matrix generated by the samples of $\tilde{h}_i(\nu)$, $i = 0, 1$ and

$$\tilde{h}_i(\nu) = \begin{cases} h_i(\nu) & \text{if } \nu \in \Theta_\beta \\ 0 & \text{if } \nu \in [0, 1] \setminus \Theta_\beta. \end{cases} \quad (42)$$

Therefore, $\tilde{h}_i(\nu)$ may be interpreted as the PSD of the measurements available at the FC under \mathcal{H}_i , $i = 0, 1$ and the test (25) can be formulated.

In the case of PFS-MAC, the covariance matrices under \mathcal{H}_0 and \mathcal{H}_1 are, respectively,

$$\begin{aligned} \Xi_{0,n'} &= \sigma_v^2 \Delta_{n'}^2 + \sigma_w^2 I_{n'}, \\ \Xi_{1,n'} &= \Delta_{n'} (F_{nn'}^H \Sigma_n F_{nn'} + \sigma_v^2 I_{n'}) \Delta_{n'} + \sigma_w^2 I_{n'}. \end{aligned} \quad (43)$$

In this case, $F_{nn'}^H \Sigma_n F_{nn'}$ is not diagonal for finite n' . However, considering that asymptotic equivalence between matrices is preserved by matrix multiplication [30, Th. 2.1 (3)], and using Lem. 1 we have that $\Sigma_n \sim F_{nn'} \text{diag}(\lambda_1^n, \dots, \lambda_{n'}^n) F_{nn'}^H$ if and only if

$$F_{nn'}^H \Sigma_n F_{nn'} \sim \text{diag}(\lambda_1^n, \dots, \lambda_{n'}^n), \quad (44)$$

which makes (41) and (43) asymptotically equivalent, obtaining the same test of hypothesis (25).

B. Proof of Theorem 4: DD Error Exponents

The asymptotic mean of the decentralized LLR statistic defined in (5) under \mathcal{H}_i , $m_i^{\text{DD}} = \lim_{n \rightarrow \infty} \mathbb{E}_i(T_d^n)$, $i = 0, 1$, depends entirely on the spectral density under \mathcal{H}_i . A similar situation occurs for the centralized detector in Th. 3. Then, using Lem. 2, we have that $m_i^{\text{DD}} = m_i(\Gamma_{\text{DD}})$ for both strategies PCS-MAC and PFS-MAC, where $m_i(\cdot)$, $i = 0, 1$, are defined in (19) and (20), respectively.

To compute the error exponents we first need to verify the assumptions of the modified version of the Gärdner-Ellis theorem with $h_0 \triangleq h_0(\nu) = \sigma_v^2 \xi(\nu) + \sigma_w^2$ and $h_1 \triangleq h_1(\nu) = (\phi(\nu) + \sigma_v^2) \xi(\nu) + \sigma_w^2$:

$$\begin{aligned} A_1) \quad & \int_{\Theta_\beta} |\log(\xi(\nu) \sigma_v^2 + \sigma_w^2)| d\nu \leq \int_{\Theta_\beta} |\log(1 + \frac{\xi(\nu) \sigma_v^2}{\sigma_w^2})| d\nu + \\ & |\log(\sigma_w^2)| \leq 1 + \frac{\sigma_v^2}{\sigma_w^2} \int_{\Theta_\beta} \xi(\nu) d\nu + |\log(\sigma_w^2)| < \infty \\ & \text{because } \xi(\nu) \geq 0, \text{ the energy constraint by (10) and } \\ & |\log(x)| < |x| \text{ if } x \geq 1. \text{ Then, } \log h_0 \in L^1([0, 1]). \\ & h_1 \in L^1([0, 1]) \text{ is proved similarly by additionally} \end{aligned}$$

considering that $\phi(\nu)$ is a power spectral density in the Wiener class.

$$A_2) \quad h_0/h_1 = 1/(1 + \Gamma_{\text{DD}}) \in L^\infty([0, 1]) \text{ given that } \text{ess inf } \Gamma_{\text{DD}} \geq 0 \text{ since } \xi(\nu) \geq 0 \text{ and } \text{ess inf } \phi(\nu) \geq 0. \\ h_1/h_0 = 1 + \Gamma_{\text{DD}} \in L^\infty([0, 1]) \text{ is easily proved by considering again that } \phi(\nu) \text{ belongs to the Wiener class and therefore } \text{ess sup } \phi(\nu) = M_\phi \text{ is a finite constant.} \\ \text{Then, } \text{ess sup } \Gamma_{\text{DD}} \leq M_\phi / \sigma_v^2.$$

Now, we apply Th. 2 to obtain the error exponents by computing the Fenchel-Legendre transforms $\Lambda_{\text{DD},i}^*(x)$ of the LMFG $\Lambda_{\text{DD},i}(t) = \log \mathbb{E}_i[e^{tT_d^n}]$, $i = 0, 1$. The same properties (P1)-(P5) in Prop. 1 are satisfied by $\Lambda_{\text{DD},i}(t)$ and $\Lambda_{\text{DD},i}^*(x)$ considering m_i^{DD} instead of m_i^{CD} , for $i = 0, 1$. Therefore, the error exponents are (26) and (27).

C. DD Optimum Energy Profile for the Miss Error Exponent

If the false alarm probability constraint is $P_{fa}^n \leq \alpha < 1$, the miss error exponent is given by Th. 4 with $\tau = m_0(\Gamma_{\text{DD}}) + \epsilon$, where $\epsilon > 0$ arbitrary small and the constraint over P_{fa}^n is satisfied for n large enough. See [33, Prop. 2] for a detailed proof. This case allows to find the optimality of the PCS-MAC and PFS-MAC strategies among all orthogonal strategies for a fixed energy profile (see [12, Th. 2]) and then, to obtain a closed form solution for the optimal energy profile $\xi(\nu)$. In this respect, we use variational calculus to solve the problem:

$$\sup \kappa_m(\xi) \text{ s.t. } \xi(\nu) \geq 0 \quad \forall \nu \in \Theta_\beta, \quad (45)$$

$$\int_{\Theta_\beta} \xi(\nu) d\nu \leq c \quad (46)$$

where $c = E_t/(\sigma_s^2 + \sigma_v^2)$. Because the error exponent $\kappa_m(\xi)$ is an increasing function of ξ (easily checked by computing the first derivative of its integrand), it is both a quasiconvex and quasiconcave (and thus quasilinear) function. The domain of the functional $\kappa_m(\xi)$ is all nonnegative functions, a convex set. The constraints of the problem are affine functions of ξ . Therefore, we have a quasiconvex optimization problem where the solution is not unique. The Lagrangian is

$$L(\xi, \lambda, \nu) = \int_{\Theta_\beta} \{-I(\xi(\nu)) - m(\nu)\xi(\nu) + \lambda(\xi(\nu) - c/\beta)\} d\nu$$

where $I(\xi(\nu)) = \frac{\xi(\nu) \sigma_v^2 + \sigma_w^2}{\xi(\nu)(\phi(\nu) + \sigma_v^2) + \sigma_w^2} + \log \frac{\xi(\nu)(\phi(\nu) + \sigma_v^2) + \sigma_w^2}{\xi(\nu) \sigma_v^2 + \sigma_w^2} - 1$ is the integrand of $\kappa_m(\xi)$. The scalar function $m(\nu)$ and the scalar constant λ are the multipliers of Lagrange.

The constraints (45) and (46) together with the following constraints are the Karush-Kuhn-Tucker necessary conditions for local extremes:

$$\begin{aligned} C_3) \quad & \lambda^* - I'(\xi^*(\nu)) \geq 0, \quad \forall \nu \in \Theta_\beta, \\ C_4) \quad & [\lambda^* - I'(\xi^*(\nu))] \xi^*(\nu) = 0, \quad \forall \nu \in \Theta_\beta, \\ C_5) \quad & \lambda^* \geq 0, \\ C_6) \quad & \lambda^* \int_{\Theta_\beta} (\xi^*(\nu) - c/\beta) d\nu = 0. \end{aligned}$$

The Euler-Lagrange equation together with the complementarity condition of $m(\nu)$ and the solution to the problem $\xi^*(\nu) \triangleq \xi_{\text{OEP}}(\nu)$ give the constraint (C4). The non-negativeness of $m(\nu)$ and λ^* produce (C3) and (C5), respectively. Finally, (C6) is due to the complementarity of λ^* and (46).

$\xi(\nu) = 0, \forall \nu$ and $\lambda = 0$ satisfies all constraints but it is not the desired solution because the error exponent is 0. Then, for a non-trivial solution $\lambda^* > 0$. It can be shown that the solutions to $I'(\hat{\xi}) = \lambda^*$ correspond to the cubic equation (29) with coefficients (30). Descartes' rule of signs of a polynomial establishes that then number of positive roots of a polynomial is related with the number of sign changes of the nonzero coefficients of consecutive powers. In the case of (30) and considering (C_5) , Descartes' rule determines that there could be at most 2 or 0 positive roots. Then, there exists a root, with a non-positive value, that does not satisfy (45) and it is discarded.

Define $\mathcal{I}_i = \{\nu \in [0, 1] : \hat{\xi}_i(\nu) > 0\}$, $i = 1, 2$ as the sets of frequencies corresponding to the positive roots. Because of Descartes' rule, $\mathcal{I}_1 \cap \mathcal{I}_2 = \mathcal{I}_1 = \mathcal{I}_2 = \Theta_{\beta^*} \subseteq \Theta_\beta$. We have two cases. If $\nu \notin \Theta_{\beta^*}$, both $\hat{\xi}_1(\nu)$ and $\hat{\xi}_2(\nu)$ are non-positive and (45) together with (C_4) imply that $\xi^*(\nu) = 0$. If $\nu \in \Theta_{\beta^*}$, both $\hat{\xi}_1(\nu)$ and $\hat{\xi}_2(\nu)$ are positive and they are possible solutions. In fact, since $I(\xi(\nu))$ is an increasing function, $\xi^*(\nu) = \max\{\hat{\xi}_1(\nu), \hat{\xi}_2(\nu)\}$. In summary, $\xi^*(\nu) = \max\{\hat{\xi}_1(\nu), \hat{\xi}_2(\nu), \hat{\xi}_3(\nu), 0\}$. A closed form for $\hat{\xi}^*(\nu)$ is shown in (31). $\lambda^* > 0$ in (C_6) implies that the energy constraint saturates, i.e., $\int_{\Theta_\beta} \xi^*(\nu) d\nu = c$ and this determines the value of λ^* . Finally, if $\nu \notin \Theta_{\beta^*}$, $\xi(\nu) = 0$ and $I(\xi(\nu)) = 0$. Then, we only need to integrate $I(\xi(\nu))$ in Θ_{β^*} to obtain the miss error exponent (28).

REFERENCES

- [1] J. Maya, L. Rey Vega, C. Galarza, and A. Altieri, "Distributed detection of correlated random processes under energy and bandwidth constraints," in *Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2014 IEEE 8th, June 2014, pp. 133–136.
- [2] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 100–117, 2012.
- [3] R. Tenney and N. R. Sandell, "Detection with distributed sensors," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-17, no. 4, pp. 501–510, July 1981.
- [4] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I: Fundamentals," in *Proc. IEEE*, vol. 85, Jan. 1997, pp. 54–63.
- [5] J. N. Tsitsiklis, "Decentralized detection," *Adv. Stat. Signal Process.*, vol. 2, pp. 297–344, Sep. 1993.
- [6] J.-F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
- [7] —, "How dense should a sensor network be for detection with correlated observations?" *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5099–5106, Nov 2006.
- [8] A. D'Costa, V. Ramachandran, and A. Sayeed, "Distributed classification of gaussian space-time sources in wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1026–1036, Aug 2004.
- [9] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1007–1015, 2004.
- [10] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *2nd Int. Symp. on Information Processing in Sensor Networks*, New York, NY, Oct. 2003.
- [11] M. Gastpar, "Uncoded transmission is exactly optimal for a simple gaussian sensor network," *Information Theory, IEEE Transactions on*, vol. 54, no. 11, pp. 5247–5251, Nov 2008.
- [12] P. Bianchi, J. Jakubowicz, and F. Roueff, "Linear precoders for the detection of a gaussian process in wireless sensors networks," *Signal Processing, IEEE Transactions on*, vol. 59, no. 3, pp. 882–894, March 2011.
- [13] K. Liu and A. Sayeed, "Type-based decentralized detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1899–1910, 2007.
- [14] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 1081–1092, March 2007.
- [15] A. Anandkumar and L. Tong, "Type-based random access for distributed detection over multiaccess fading channels," *Signal Processing, IEEE Transactions on*, vol. 55, pp. 5032–5043, 2007.
- [16] J. Maya, L. Rey Vega, and C. Galarza, "Error exponents for bias detection of a correlated process over a mac fading channel," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2013 IEEE 5th International Workshop on, Dec 2013, pp. 484–487.
- [17] W. Li and H. Dai, "Distributed Detection in Wireless Sensor Networks Using A Multiple Access Channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, 2007.
- [18] S. Dasarathan and C. Tepedelenioglu, "Distributed Estimation and Detection With Bounded Transmissions Over Gaussian Multiple Access Channels," vol. 62, no. 13, pp. 3454–3463, 2014.
- [19] K. Cohen and A. Leshem, "Performance analysis of likelihood-based multiple access for detection over fading channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2471–2481, 2013.
- [20] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent gaussian noise," *Signal Processing, IEEE Transactions on*, vol. 48, no. 12, pp. 3266–3279, 2000.
- [21] Y. Sung, H. Poor, and H. Yu, "How much information can one get from a wireless ad hoc sensor network over a correlated random field?" *Information Theory, IEEE Transactions on*, vol. 55, no. 6, pp. 2827–2847, June 2009.
- [22] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "Cooperative routing for distributed detection in large sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 2, pp. 471–483, February 2007.
- [23] A. Anandkumar, L. Tong, and A. Swami, "Detection of gauss-markov random fields with nearest-neighbor dependency," *Information Theory, IEEE Transactions on*, vol. 55, no. 2, pp. 816–827, Feb 2009.
- [24] C. Tepedelenioglu and A. Narasimhamurthy, "Universal distributed estimation over multiple access channels with constant modulus signaling," *Signal Processing, IEEE Transactions on*, vol. 58, no. 9, pp. 4783–4794, Sept 2010.
- [25] M. Banavar, C. Tepedelenioglu, and A. Spanias, "Distributed snr estimation with power constrained signaling over gaussian multiple-access channels," *Signal Processing, IEEE Transactions on*, vol. 60, no. 6, pp. 3289–3294, June 2012.
- [26] S. Dasarathan and C. Tepedelenioglu, "Distributed estimation and detection with bounded transmissions over gaussian multiple access channels," *Signal Processing, IEEE Transactions on*, vol. 62, no. 13, pp. 3454–3463, July 2014.
- [27] M. Goldenbaum and S. Stanczak, "Computing the geometric mean over multiple-access channels: Error analysis and comparisons," in *Signals, Systems and Computers (ASIOMAR)*, 2010 Conference Record of the Forty Fourth Asilomar Conference on, Nov 2010, pp. 2172–2178.
- [28] F. Li and J. Evans, "Optimal strategies for distributed detection over multiaccess channels," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, March 2008, pp. 2417–2420.
- [29] S. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, 1st ed. Prentice-Hall, 1998.
- [30] R. M. Gray, "Toeplitz and circulant matrices: A review," Tech. Rep., 2001.
- [31] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. Springer, 1998.
- [32] B. Bercu, F. Gamboa, and A. Rouault, "Large deviations for quadratic forms of stationary Gaussian processes," *Stochastic Processes and their Applications*, vol. 71, no. 1, pp. 75–90, 1997.
- [33] W. Zhang, H. Poor, and Z. Quan, "Frequency-domain correlation: An asymptotically optimum approximation of quadratic likelihood ratio detectors," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 969–979, March 2010.

- [34] P. Tilli, "Fast reliable algorithms for matrices with structure," T. Kailath and A. H. Sayed, Eds. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1999, ch. Asymptotic Spectral Distribution of Toeplitz-related Matrices, pp. 153–187.